Noninvasive neurostimulation of left ventral motor cortex enhances sensorimotor adaptation in speech production

Noninvasive neurostimulation of left ventral motor cortex enhances sensorimotor adaptation in speech production

Terri L. Scott[1,2], Laura Haenchen[1], Ayoub Daliri[1,3], Julia Chartove[2], Frank H. Guenther[1,4], & Tyler K. Perrachione[1,†]

[1]Department of Speech, Language, & Hearing Sciences, Boston University, Boston, MA 02215
[2]Graduate Program for Neuroscience, Boston University, Boston, MA 02215
[3]College of Health Solutions, Arizona State University, Tempe, AZ 85287
[4]Department of Biomedical Engineering, Boston University, Boston, MA 02215

†Correspondence:
Tyler K. Perrachione, Ph.D.
Department of Speech, Language, and Hearing Sciences, Boston University
635 Commonwealth Ave.
Boston, MA 02215
Phone: +1.617.358.7410
Email: tkp@bu.edu

Abstract

Sensorimotor adaptation—enduring changes to motor commands due to sensory feedback—allows speakers to match their articulations to intended speech acoustics. How the brain integrates auditory feedback to modify speech motor commands and what limits the degree of these modifications remain unknown. Here, we investigated the role of speech motor cortex in modifying stored speech motor plans. In a within-subjects design, participants underwent separate sessions of sham and anodal transcranial direct current stimulation (tDCS) over speech motor cortex while speaking and receiving altered auditory feedback of the first formant. Anodal tDCS increased the rate of sensorimotor adaptation for feedback perturbation. Computational modeling of our results using the Directions Into Velocities of Articulators (DIVA) framework of speech production suggested that tDCS primarily affected behavior by increasing the feedforward learning rate. This study demonstrates how focal noninvasive neurostimulation can enhance the integration of auditory feedback into speech motor plans.

Keywords
sensorimotor adaptation; speech production; tDCS; premotor cortex; auditory feedback control; auditory feedback perturbation; formant perturbation; motor learning;

1. Introduction

The brain maintains fast and precise motor actions by adapting learned motor commands to changing conditions. When there is a sustained mismatch between intended motor events and sensory feedback, the motor system exhibits *sensorimotor adaptation*: feedback-based motor learning that accumulates over longer timescales to change established motor plans. Sensorimotor adaptation serves a particularly important role during speech production, where acoustics of contrastive speech categories depend on minuscule articulation differences, yet speech intelligibility must be preserved during vocal tract ontogeny (Redford, 2019). In order to execute the specialized motor actions required for accurate speech under varying conditions, the speech motor system must be able to incorporate information from sensory feedback into established feedforward motor commands (Guenther, 2016). This ability to use sensory feedback to alter feedforward commands accounts for how consistent speech execution is maintained under physical changes to the vocal tract, such as in typical development and aging, and injury.

Sensorimotor adaptation has been demonstrated experimentally for several different auditory characteristics of speech using feedback perturbations (Houde and Jordan, 1998; Jones and Munhall, 2000; Purcell and Munhall, 2006; Villacorta et al., 2007; Shiller et al., 2009). Speakers' compensatory response typically adjusts their speech productions to oppose the perceived acoustic perturbation; however, the cortical mechanisms that support the integration of auditory feedback with motor planning are unknown. Speech motor control models, such as the Directions Into Velocities of Articulators (DIVA) model (Guenther, 1994, 1995; Guenther et al., 2006; Golfinopoulos et al., 2010), posit that motor programs for common phoneme sequences are represented in left ventral premotor cortex (vPMC) and serve as templates against which to compare sensory feedback during speech production. Mismatch between these learned motor representations and auditory feedback is thought to be transformed into compensatory gestures in ventral motor cortex (vMC; Tourville and Guenther, 2011). Correlational support for this model comes from neuroimaging studies in which neural activation in these regions is found during speech production (Ghosh et al., 2008; Tourville et al., 2008; Basilakos et al., 2018) and is proportional to speakers' compensation for unexpected, intermittent, auditory feedback perturbations (Niziolek and Guenther, 2013; Behroozmand et al., 2015).

The first aim of our study was to determine the effect of noninvasive neurostimulation applied to left ventral sensorimotor cortex on sensorimotor adaptation to auditory perturbation of speech. Participants underwent an established speech sensorimotor adaptation protocol with perturbed auditory feedback while we measured the magnitude and rate of sensorimotor adaptation reflected by changing speech acoustics. To modulate neural function of left ventral sensorimotor cortex during the task, participants simultaneously received transcranial direct current stimulation (tDCS)—a noninvasive neurostimulation technique in which a low current is applied over the scalp via electrodes to induce small changes to the electric field in underlying cortex. The polarity of current flow is believed to determine the effect of stimulation on cortical function, with anodal stimulation increasing neural excitability and cathodal stimulation decreasing excitability (Nitsche and Paulus, 2000; Dayan et al., 2013; Filmer et al., 2014). Additionally, tDCS is believed to modulate cortical plasticity, as its neuromodulatory effects can be measured for some time after stimulation has ceased (Nitsche and Paulus, 2000; Nitsche et al., 2007; Rroji et al., 2015). In the language domain, tDCS has been demonstrated to facilitate word naming (Fertonani et al., 2010; Malyutina and den Ouden, 2015) and production of difficult phoneme sequences (Buchwald et al., 2018), amongst other language tasks (reviewed in Monti et al., 2013).

We found that anodal tDCS of left sensorimotor cortex was indeed associated with increase in the rate of sensorimotor adaptation to speech. However, because multiple neural mechanisms may be affected by tDCS, the second aim of this study was to ascertain, in mechanistic terms, how tDCS may have affected cortical function for speech motor adaptation using computational simulations of the DIVA model. We identified several candidate neurocomputational variables that could hypothetically be altered by anodal

tDCS in our experiment. First, by increasing cortical excitation under tDCS, increased sensitivity to auditory errors could elicit a greater compensatory response associated with within-trial auditory feedback control-based mechanisms. Alternatively, tDCS could act to modulate trial-to-trial adaptation (plasticity) and increase learning-based anticipatory corrections. Ultimately, computational modeling favored an effect of tDCS on the rate of trial-to-trial adaptation of motor programs, as well as a small decrease in sensitivity to somatosensory feedback that normally opposes compensation to perturbed auditory feedback (Nasir and Ostry, 2009; Katseff et al., 2011; Lametti et al., 2012). These results expand our understanding of the neurobiological bases of speech adaptation. Furthermore, these results demonstrate the explanatory power of combining neurostimulation with computational modeling to make inferences about cortical function.

2. Materials and Methods

2.1 Participants

Right-handed, native speakers of American English free from speech, language, or hearing deficits completed this study ($N = 18$; 4 male, 14 female; age 18-28 years, $M = 20.4 \pm 2.1$). Because not all speakers adapt to auditory perturbation (Purcell and Munhall, 2006; Lametti et al., 2012), we recruited a total of 37 participants (10 male, 27 female; age 18-28 years, $M = 21.3 \pm 2.2$) to complete an initial screening session that did not involve tDCS. To be included in the tDCS portion of the study, we required each participant to demonstrate sensorimotor adaptation such that auditory perturbation resulted in significantly lower first formant (F1) frequencies relative to his or her own baseline productions (see §2.2.1 below). We used a two-sample, one-sided Kolmogorov-Smirnov test to test whether each participant exhibited significant ($p < 0.001$) adaptation during the second half of the perturbation phase of the experiment relative to the baseline trials. Twenty-three recruited participants met our inclusion criterion (5 male, 18 female; 18-28 years, $M = 21.08 \pm 2.4$), but five of these did not complete one or both tDCS sessions and were withdrawn from the study. Of the 14 participants who did not meet our inclusion criteria for sensorimotor adaptation, two exhibited "following" responses (Burnett et al., 1998), as determined by another Kolmogorov-Smirnov testing whether F1 frequencies measured during the second half of the perturbation phase were significantly higher than at baseline. Screening session data for all recruited participants is given in supplemental **Figure S1** Participants provided written informed consent, approved and overseen by the Institutional Review Board at Boston University, and were paid for their participation.

2.2 Experimental Design

Participants completed three sessions in which they underwent the same behavioral paradigm (**Figure 1A**). Each session was separated from the previous by at least 7 days to reduce the potential for carry-over of learning across sessions. In an initial session without tDCS, we confirmed that participants adapted to auditory feedback perturbation of their speech. Participants were then assigned to receive either anodal or sham stimulation during their second session and the other during their third session. Nine participants completed each order of stimulation. Although participants were told before each tDCS session that they would either be receiving active stimulation or be in a control condition (sham), participants were blind not only to which condition they were in on each visit, but also to the fact that they would be in both conditions on counterbalanced visits.

2.2.1 Behavioral paradigm

Each session was conducted in a sound-attenuated chamber. Stimulus delivery, recording, and real-time resynthesis for auditory perturbation were controlled via the Audapter software (Cai et al., 2008)

implemented in MATLAB vR2014b (The Mathworks, Natick, MA). Participants' speech was transduced using a Shure MX153 earset microphone, Behringer Ultragain Pro two-channel microphone amplifier, and Roland Quad Capture sound card. Auditory stimulation was delivered via the same sound card, an Art HeadAmp6 Pro headphone amplifier, and Etymotic ER-3C insert earphones.

Participants were prompted by the Audapter software to say the words "bed," "dead," and "head," in a pseudorandom order. The paradigm began with a brief training phase, in which participants received feedback to insure they were producing the words at a sufficient loudness (72–88 dB SPL) and duration (400–600 ms); trials in the training phase were repeated until productions of suitable intensity and duration were achieved. Participants continued to receive feedback about the intensity and duration of their speech during the experiment, but trials were not repeated. The Audapter software performed real-time analysis, replay, resynthesis, and recording of participants' speech acoustics (F1 and F2) (**Figure 1B**).

The behavioral paradigm consisted of four phases (**Figure 1C**). The *baseline phase* consisted of 57 trials in which participants spoke the target words and heard their own, unperturbed speech as auditory feedback. Next, during the *ramp phase*, real-time perturbation of participants' F1 was introduced at +15% for 3 trials; the ramp phase was included to reduce participants' conscious detection of the auditory perturbation, but was kept brief to allow us to observe continued learning during the subsequent shift phase. During the *perturbation phase*, which lasted for 60 trials, participants heard as auditory feedback a real-time perturbation of their own speech in which F1 was increased by 30%. Finally, during the 60 trials of the *return phase*, participants again heard their own, unperturbed speech as auditory feedback. Auditory feedback was presented at 5 dB SPL above the participant's own productions.

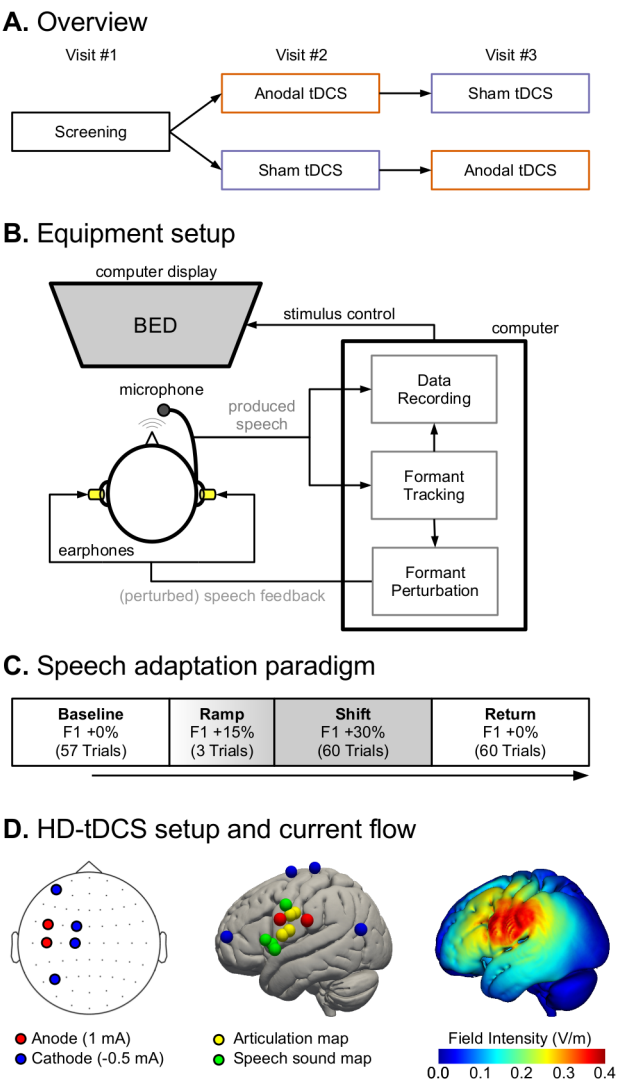

Figure 1: Paradigm design and tDCS stimulation
(A) During their first visit to the lab, participants underwent an initial session of the experiment without tDCS to confirm they adapted to auditory perturbations. Over two subsequent visits, they completed the tDCS sessions, with order counterbalanced across participants. **(B)** Schematic of the equipment setup and behavioral paradigm. **(C)** The behavioral paradigm was the same on each visit. Participants' baseline speech acoustics was measured without perturbation; then F1 perturbation was increased to +15% during the *ramp* phase, held at +30% during the *shift* phase, and presented again without perturbation during the *return* phase. **(D)** Location of anodes (FC5, C5; red points) and cathodes (AF7, FC1, C1, P5; blue points) in the 10-10 electrode system (left) and with their location overlaid on the cortical surface (middle). Also shown are the locations of the articulator maps (tongue and jaw; yellow points) and speech sound maps (green points) from the DIVA model (Table D.1 in Guenther, 2016, p391). Estimated cortical surface field intensity from this stimulation montage is shown at right.

2.2.2 tDCS stimulation

Neurostimulation was controlled and delivered using a Soterix MxN high-definition (HD) tDCS system. HD-tDCS was used both because it offers more focused stimulation and avoids strong effects of equal and opposite current density in brain areas outside of the region of interest. Stimulating electrodes (2 mA) were placed at FC5 and C5, and return electrodes were placed at AF7, FC1, C1, and P5, in a roughly center-surround configuration (Datta et al., 2009; Kuo et al., 2013). This montage was selected to optimize field intensity and current flow over left vPMC and vMC (**Figure 1D**), as determined by simulation using the HD-Explore software (Soterix Medical Inc.; Datta et al., 2009; Huang et al., 2017). These areas were targeted in this study because they are the theoretical location of the speech sound maps and articulator maps for feedforward control of speech production (Tourville and Guenther, 2011; Guenther, 2016).

After insuring the resistance of each channel was < 10 kΩ, anodal stimulation began with a 30-s linear ramp from 0 to 2 mA, with tonic 2 mA stimulation continuing for the remainder of the session (~20 min). The procedure for sham stimulation was the same, but after the 30-s ramp to 2 mA, stimulation was linearly decreased over 30 s back to 0 mA, where it remained throughout the behavioral paradigm. This procedure effectively blinded participants to whether they were receiving anodal or sham stimulation during the behavioral task, which was begun 60 s after the onset of stimulation (see §3.1, below). Stimulation began before the training phase and ended after the last trial of the return phase for a mean duration of 17 min 13 s (range: 16 min 36 s - 18 min 8 s). The mean durations of the different phases of the experiment (excluding the screening session) are as follows: training, 1 min 46 s; baseline, 4 min 54 s; ramp, 15 s; full perturbation, 5 min 9 s; and return, 5 min 3 s.

2.3 Statistical Analysis

Speech acoustics (mean F1 and F2 frequencies) were obtained from each participant on each trial in each condition using Audapter. Vowel formant frequencies were isolated by analyzing 60% of the word's duration beginning 10% after the onset of voicing. Outlier trials, in which F1 deviated by more than two standard deviations from the mean value in the respective session and phase, were excluded from the analysis (< 5% of total trials). In a repeated-measures analysis of variance (ANOVA), the number of F1 outliers did not differ as function of *stimulation* (no-tDCS, anodal, or sham; $F_{2,34} = 0.29$, $p = 0.75$), *phase* (baseline, perturbation, return; $F_{2,34} = 0.16$, $p = 0.85$), or their interaction ($F_{4,68} = 1.12$, $p = 0.36$). Participants' F1 and F2 measurements were then normalized (proportionally) to the mean F1 and F2 values obtained during the baseline phase of each session. To control for errors in production and automated formant tracking errors, spectrograms of all trials were visually inspected using the Praat software (Boersma, 2001), and F1 and F2 were measured manually and compared to the program's measured values to insure accurate formant measures on each trial.

Speech acoustics data were analyzed in R using linear mixed-effects models implemented in the package *lme4* (Bates et al., 2014). The models' fixed-effect terms included categorical factors for *stimulation* (anodal vs. sham) and *session* (2 vs. 3), the mean-centered continuous factor *time* (trial), and the *stimulation* × *time*, *stimulation* × *session*, *session* × *time*, and *stimulation* × *time* × *session* interactions. The models' random effects terms included by-participant intercepts, by-participant slopes for the fixed factors *stimulation*, *time*, and *session* and by-item intercepts for each word. Statistical comparisons of model terms were determined via application of deviation-coded contrasts to the model matrix. Significance of main effects and interactions was determined by adopting a significance criterion of $\alpha = 0.05$, with *p*-values for model terms based on the Satterthwaite approximation of the degrees of freedom obtained from the package *lmerTest* (Kuznetsova et al., 2017).

Additional analysis using nonlinear (exponential) models was conducted; however, while the group average perturbation curve is exponential (**Figure 2A**), in many cases individual participants' adaptation during the perturbation phase was not well described by an exponential function (e.g., when participants exhibited no adaptation during the perturbation phase in some condition; see **Figure 3B** for individual data). For participants and stimulation conditions where adaptation was evident, linear and exponential adaptation models did not differ in fit. Correspondingly, we chose to model these results using linear mixed effects models because of their power and precision. Further, we chose *ab initio* to employ models with maximal fixed and random effects structures, (Barr et al., 2013), as the purpose of these models was confirmatory hypothesis testing rather than model selection (Meteyard & Davies, 2020). For analyses with only a single value per participant per factor level, data were analyzed using repeated-measures ANOVA in the package *ez* (Lawrence, 2013).

3. Results

3.1 Somatic and psychological experiences related to tDCS

After each tDCS session, participants completed a questionnaire detailing the presence and severity of any symptoms or side effects they experienced, as well as whether they believed these effects to be related to the administration of tDCS (Bruononi et al., 2011). Participants filled out identical forms after both sessions, as they were not told whether they had received anodal or sham stimulation in each session. Nearly all participants reported mild to moderate tingling sensations in both anodal and sham sessions. Less frequently, participants reported experiencing pain or burning on their scalp. The prevalence or intensity of these sensations did not differ between the anodal and sham conditions, suggesting that participants were effectively blinded to whether they were receiving active or control stimulation (**Figure S2**). We did not inquire directly as to whether participants thought they had received sham or anodal stimulation. Several participants also reported feeling sleepy or distracted, but unlike their somatic experiences, participants rarely attributed their state of arousal to tDCS.

3.2 Primary outcome measures: adaptation and recovery

3.2.1 Speech adaptation during perturbation

We determined whether participants' motor adaptation to auditory F1 perturbation during speech production was affected by tDCS in a linear mixed-effects model of trial-by-trial F1 adaptation magnitude (% of mean baseline F1) during the perturbation phases of the anodal and sham tDCS conditions. A corresponding model was run on F2 acoustics as a control, as auditory feedback of F2 was not perturbed. We omitted the "ramp" phase of the perturbation from these analyses and focused only on the trials for which the feedback perturbation remained constant.

We observed a significant main effect of *time*, such that adaptation (the lowering of F1) increased over the perturbation period (β = -0.0005, *s.e.* = 0.0001, t = -4.65, p = 0.0003) in all conditions. Importantly, there was a significant *stimulation* × *time* interaction such that participants showed greater adaptation with time under anodal stimulation than sham (**Figure 2A**; β = 0.0002, *s.e.* = 0.0001, t = 2.99, p = 0.003). We also observed a significant *time* × *session* interaction such that participants showed a greater rate of adaptation during session 2 vs. session 3 (**Figure 3A**; β = -0.0001, *s.e.* = 0.0001, t = -2.40, p = 0.02). We did not observe a main effect of *stimulation* (β = 0.0061, *s.e.* = 0.0058, t = 1.06, p = 0.31) when considering the entire perturbation period. No significant effect of *session*, no *stimulation* × *session* interaction, and no three-way interaction was observed (see **Table S1**). Speakers' F1 during the latter half of the perturbation phase under anodal stimulation was 91.4% ± 4.5% that of the baseline, whereas under sham stimulation it was 93.1 ± 5.0% of baseline, corresponding to compensation of 28.7% and 23.0% of the full

auditory perturbation, respectively (**Figure 2B**). Individual participant data for sham and anodal tDCS sessions are given in **Figure 3B**.

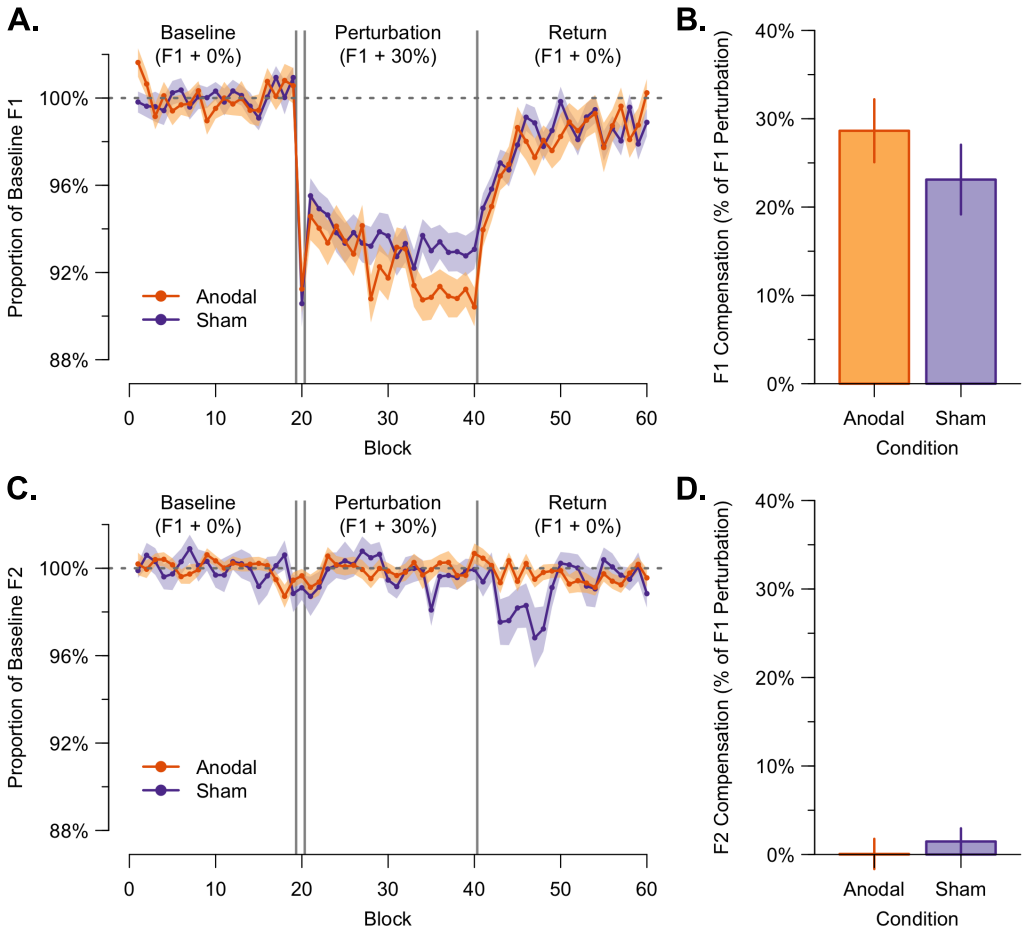

Figure 2: Speech adaptation under perturbation during anodal tDCS vs. sham.
(A) In response to a perceived increase in F1 frequency during the auditory feedback perturbation phase of each session, participants compensated by lowering the F1 frequency of their own speech productions. The rate of adaptation under anodal tDCS (orange) was significantly enhanced relative to sham tDCS (purple). Blocks represent averages across 3 trials (consisting of 1 trial for each of the three presented words) calculated within participants. Shaded regions indicate standard error of the mean across participants for each block. Vertical lines indicate the onset/offset of each phase, with the two lines before the perturbation phase indicating the brief ramp phase. **(B)** The average magnitude of compensatory responses scaled with respect to the full perturbation (+30% of baseline) during the latter half of the perturbation phase of the anodal (orange) and sham tDCS sessions (purple). Error bars represent the standard error of the mean across participants. **(C)** We measured speakers' F2 values to test if sensorimotor adaptation under tDCS was confined to F1. We did not observe systematic changes to F2 under anodal tDCS (orange) or sham tDCS (purple) during the perturbation. Shaded regions indicate standard error of the mean across participants for each block. **(D)** The average magnitude of F2 relative to baseline and scaled by the same percent factor as F1 in **(B)** during the latter half of the perturbation phase of the anodal (orange) and sham tDCS sessions (purple). Error bars represent the standard error of the mean across participants.

The corresponding model of F2 showed no effects of *stimulation* (**Figure 2C**; β = -0.0007, *s.e.* = 0.0027, t = -0.26, p = 0.80), *time* ($\beta = 1.6 \times 10^{-5}$, *s.e.* = 0.0001, t = 0.22, p = 0.83), *session* (**Figure 3A**; β = 0.0041, *s.e.* = 0.0027, t = 1.53, p = 0.15) on this unperturbed feature, and no significant interactions (**Table S2**). Mean F2 values during the latter half of the perturbation phase were 100.0% ± 2.2% of baseline during anodal tDCS and 99.6% ± 2.0% during sham (**Figure 2D**).

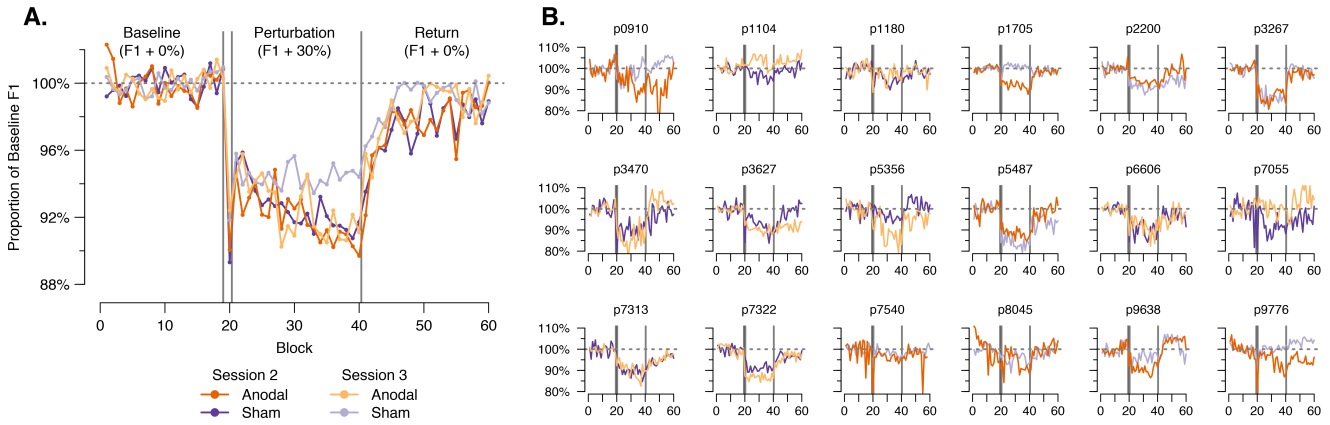

Figure 3. Sensorimotor adaptation by stimulation condition, session, and participant.
(A) Lines depict the average F1 percent change from baseline across participants. Nine participants contribute to each line. Anodal tDCS is shown in orange, while sham tDCS is shown in purple, with darker colors representing Session 2 and lighter colors representing Session 3. Note that within-subject comparisons are between lines of opposite color and shading (e.g., Session 2 Anodal with Session 3 Sham). The double vertical gray line indicates the beginning and end of the "ramp" phase. **(B)** Individual participants sensorimotor adaptation during anodal and sham tDCS sessions. Each panel corresponds to a single participants' data; participant numbers (e.g., "p0910") were assigned randomly and are not sequential. All participants are shown on the same scale; F1 difference from baseline was averaged over blocks of three trials for display. The double vertical gray line indicates the beginning and end of the "ramp" phase. Color and shading conventions as in panel (A).

3.2.2 Speech recovery following perturbation

We also analyzed whether participants' motor recovery after removal of auditory perturbation of F1 was affected by tDCS using a linear mixed-effects model of trial-by-trial F1 difference from baseline during the return phases of the anodal and sham tDCS conditions. This model included the same fixed and random factors described in §3.2.1 and modeled trials beginning with the last trial in which the auditory perturbation was presented through the end of the session. The *time* factor was centered on its mean value.

We observed a significant effect of *time* as participants' F1 values gradually returned to baseline (**Figure 2A**; $\beta = 0.0006$, *s.e.* $= 0.0001$, $t = 5.82$, $p \ll 0.0001$). We did not observe significant effects of *stimulation* or *session*, nor any significant interactions between factors (**Table S3**). We ran the corresponding model on the participants' F2 values during the return phase and found significant interactions between *stimulation* and *time* (**Figure 2C**; $\beta = 0.0002$, *s.e.* $= 0.0001$, $t = 3.65$, $p = 0.0002$) and *stimulation* and *session* ($\beta = -0.0086$, *s.e.* $= 0.0027$, $t = -3.13$, $p = 0.006$). (These results appear to have been driven primarily by one participant who exhibited substantial inconsistency in their F2 productions during their sham session; whereas most participants showed no effect of stimulation, perturbation, or its withdrawal on F2 productions.)

3.3 Secondary outcomes and control measures

3.3.1 Consistency in individual differences in compensation across conditions

We investigated whether the average magnitude of adaptation during the latter half of the perturbation phase (where production targets were expected to most approach stability) was consistent within participants across the stimulation conditions using Spearman's rank correlation. Participants' F1 adaptation in the initial session without tDCS was not significantly correlated with their adaptation during sham stimulation ($r = 0.41$, $p = 0.09$) or during anodal stimulation ($r = 0.39$, $p = 0.11$). The magnitude of adaptation was also not significantly correlated between sham and anodal stimulation conditions ($r = 0.42$, $p = 0.09$).

Additionally, we investigated whether participants' magnitude of adaptation was correlated across sessions regardless of stimulation condition. F1 adaptation was significantly correlated between the screening session and the second session ($r = 0.55$, $p = 0.02$), while the correlation between screening and the third session was not significant ($r = 0.45$, $p = 0.06$). We ran a *post-hoc* analysis to better understand whether these two last correlations were significantly different from each other due to the fact that they fell on either side of our significance criterion, and found that they were not (Pearson and Filon's $z = 0.48$, $p = 0.31$, implemented in the R package *cocor*; Diedenhofen & Musch, 2015).

3.3.2 Speech production variability under tDCS

We also investigated whether the coefficient of variation ($s/\bar{x}$; a measure of instability obtained from speech variability across individual trials) for participants' F1 differed as a function of *stimulation* (anodal, sham stimulation) during each *phase* of the experiment (baseline, perturbation, return). We limited analysis to the latter half of the perturbation and return phases to avoid biased coefficients of variation resulting from effects related to the initial administration and cessation of auditory perturbation. In a repeated-measures ANOVA of the coefficient of variation of F1 with within-subject factors of *stimulation* and *phase*, we found a significant effect of *stimulation* ($F_{1,17} = 5.17$, $p = 0.04$, $\eta^2_G = 0.05$) such that the coefficient of variation tended to be greater under anodal than sham stimulation, no effect of *phase* ($F_{2,34} = 1.55$, $p = 0.23$, $\eta^2_G = 0.02$), and no *stimulation* × *phase* interaction ($F_{2,34} = 0.52$, $p = 0.60$, $\eta^2_G = 0.01$). This analysis was repeated for participants' coefficient of variation of F2 in which we found no significant effects of *stimulation* ($F_{1,17} = 1.42$, $p = 0.25$, $\eta^2_G = 0.01$), *phase* ($F_{2,34} = 0.59$, $p = 0.56$, $\eta^2_G = 7.3 \times 10^{-4}$), and no *stimulation* × *phase* interaction ($F_{2,34} = 0.32$, $p = 0.72$, $\eta^2_G = 5.6 \times 10^{-4}$).

3.3.3 Speech production baseline under tDCS

To determine whether the application of tDCS had an effect on speech production acoustics independent of the perturbation manipulation, we performed a series of linear-mixed effects models testing whether speakers' F1 frequency during the baseline phase was affected by *session* (1, 2, or 3) and *stimulation* (no-tDCS, anodal, or sham).

We first tested a linear mixed-effects model including a categorical fixed factor for all levels of *session*, random slopes and intercepts by participant, and random item intercepts. (The *stimulation* factor was not included in this model because, for all participants, the first session did not involve any tDCS, meaning these levels of the two factors were perfectly colinear, and a model including both together would be rank deficient.) We found a significant difference in baseline F1 productions between the two earlier sessions (Session 2 – Session 1; $\beta = -10.67$, *s.e.* = 4.20, $t = -2.54$, $p = 0.02$), whereas we did not find any difference in the latter two sessions (Session 3 – Session 2; $\beta = -4.12$, *s.e.* = 8.28, $t = -0.50$, $p = 0.63$). F1 values across participants were highest during the first visit (701 ± 95 Hz), lower during the second visit (690 ± 97 Hz), and lowest during the third visit (686 ± 97 Hz). The direction of this learning effect (lowered F1 values across sessions) is consistent with long-term retention of adaptation for the auditory perturbation (raised F1 values) in this study.

Additionally, we tested whether the order of stimulation sessions affected baseline F1 frequency; e.g., if receiving anodal stimulation during Session 2 was associated with a greater change in F1 baseline at Session 3. In an ANOVA on a second model including only the second and third session baselines, with categorical fixed factors including *session* (2 vs. 3), *stimulation* (anodal vs. sham), and their interaction, and random factors including by-participant intercepts and by-participant slopes for the effect of *stimulation*, we found no effect of *session* ($F_{1,16} = 0.22$, $p = 0.65$), *stimulation* ($F_{1,16} = 0.0047$, $p = 0.84$) and no *stimulation* by *session* interaction ($F_{1,16} = 0.0013$, $p = 0.97$), suggesting the tDCS manipulation did not affect learning across sessions.

3.3.4 Amount of compensation to initial perturbation trial

The primary outcome measures indicated that anodal tDCS had an effect on the rate of adaptation during the perturbation phase; however, the prior literature distinguishes *compensatory* (or reflexive) responses to unexpected perturbation from *adaptive* responses to ongoing sensorimotor mismatch that involves modifications to feedforward commands (Burnett et al., 1998; Guenther, 2016). We therefore also investigated whether reflexive response to the initial application of perturbed auditory feedback on the first trial of the ramp phase differed between conditions. In a repeated measures ANOVA of F1 compensation on the first perturbation trial, with *stimulation* (anodal vs. sham) as the within-subjects factor, we found no effect of tDCS on the response magnitude to initial perturbation ($F_{1,17} = 0.04$, $p = 0.84$, $\eta^2_G = 0.002$; anodal: 94.0% ± 8.3%; sham: 94.5% ± 5.3%). In a corresponding analysis of *session* (1 vs. 2 vs. 3) we observed a trend for the magnitude of the reflexive response on the first perturbation trial to decrease as a function of experience with the task (Session 1: 91.8% ± 7.3% of baseline; Session 2: 92.9% ± 7.9%; Session 3: 95.7% ± 5.6%); however, this trend was not statistically significant ($F_{1,17} = 3.62$, $p = 0.07$, $\eta^2_G = 0.18$). Finally, reflexive response magnitude to the initial perturbation trial was not correlated with speakers' overall adaptation magnitude, either during the initial visit ($r = -0.19$, $p = 0.44$), anodal tDCS ($r = 0.11$, $p = 0.67$), or sham tDCS ($r = 0.22$, $p = 0.36$).

4. Computational Modeling

4.1 Model description

Several distinct motor control mechanisms can contribute to compensatory responses during motor adaptation under sensory perturbations (Scott, 2004; Shadmehr et al., 2010). To investigate which aspects of motor learning and performance were responsible for changes in adaptive responses under neurostimulation, we performed computer simulations using SimpleDIVA (Kearney et al., 2020)—a simplified version of the DIVA model (Guenther et al., 2006; Guenther, 2016) that characterizes the neural computations involved in speech motor control. The SimpleDIVA model is designed to capture the aggregate contributions of DIVA model's auditory feedback control, somatosensory feedback control, and feedforward control subsystems to speech acoustics, without needing to model each system's various components in detail. Further, rather than modeling the configuration and trajectory of the various vocal tract articulators in detail, SimpleDIVA abstracts motor control to the realized acoustic output, here F1 frequency. In this implementation, SimpleDIVA accounts for trial-by-trial changes in speech acoustics by estimating the aggregate contributions of these three subsystems to speech acoustics during sensorimotor adaptation experiments.

The first mechanism that contributes to compensatory responses is the auditory feedback control subsystem of the speech motor controller. This subsystem translates production errors detected via the auditory system into corrective movements, with a latency of approximately 100-200 ms from error/perturbation onset to the start of the corrective movement. We will refer to this within-trial component of the compensatory response as the *reflexive response*, borrowing terminology from Larson and colleagues (Burnett et al., 1998; Hain et al., 2000), while noting that this "reflex" involves processing in the cerebral cortex (Tourville et al., 2008). The term *auditory feedback control gain* (α_A) will be used to describe the size of this response relative to the size of the auditory error. A gain of 1 would indicate that the auditory feedback control system is completely counteracting the perturbation, but in actuality the auditory feedback control gain appears to be much smaller, with prior studies indicating a compensatory response that is typically less than 25% of the size of the perturbation (e.g., Burnett et al., 1998; Chen et al., 2007; Tourville et al., 2008; Niziolek and Guenther, 2013). Compensatory responses to auditory perturbations have the effect of generating somatosensory feedback that no longer matches the motor system's expectations (somatosensory target) for the speech gesture. This will invoke somatosensory

feedback control mechanisms that tend to counteract the compensatory response. The size of the somatosensory feedback controller's opposition to the compensatory response will depend on the *somatosensory feedback control gain* (α_S).

If the perturbation is sustained over many productions of the same sound, a second mechanism is invoked by the motor system to counteract the perturbation: trial-to-trial adaptation of the feedforward command, or stored "motor program." We will refer to this as the *adaptive response*, the size of which is modulated by the *feedforward command learning rate* (λ_{FF}). Thus, in the terminology used here, the compensatory response to a sustained perturbation is composed of a reflexive response and an adaptive response.

The following equations used in the current simulations capture the key aspects of the DIVA model in a simplified form (Kearney et al., 2020) that involves only three free parameters (α_A, α_S, and λ_{FF}), thereby eliminating redundancies in the set of fitting parameters that would otherwise obfuscate the neural mechanisms underlying compensation since such redundancies can lead to multiple parameter values that produce equivalent fits to the data. **Equation 1** defines the value of F1 produced by the subject on a given trial (indexed by n) as:

$$F1_{produced}(n) = F1_{FF}(n) + \Delta F1_{FB}(n) \quad \text{(EQ 1)}$$

In words, the F1 value produced on a trial is a combination of a feedforward command ($F1_{FF}$) and a sensory feedback-based correction ($\Delta F1_{FB}$) that is initiated if/when the auditory and somatosensory feedback controllers detect production errors on the current trial. At the start of each simulation, $F1_{FF}$ is initialized to the average F1 measured during the baseline phase of the experiment across participants. **Equation 2** defines the feedback-based correction as:

$$\Delta F1_{FB}(n) = \alpha_A \times (F1_{AT} - F1_{perceived}(n)) + \alpha_S \times (F1_{ST} - F1_{FF}(n)) \quad \text{(EQ 2)}$$

where $F1_{AT}$ and $F1_{ST}$ are the F1 values specified by previously learned auditory and somatosensory targets, respectively, for the vowel; $F1_{perceived}$ is the value of F1 heard by the subject (including the perturbation, when one is applied) before feedback control mechanisms kick in on that trial (i.e., $F1_{perceived} = F1_{FF}(n)$ + perturbation size); and α_A and α_S are the gains of the auditory and somatosensory feedback control subsystems, respectively. In the simulations, $F1_{AT}$ and $F1_{ST}$ are set to the average F1 of the baseline phase, corresponding to the assumption that the auditory and somatosensory targets will not change substantially over the course of the experiment. **Equation 3** describes the procedure for updating the feedforward command from trial to trial:

$$F1_{FF}(n+1) = F1_{FF}(n) + \lambda_{FF} \times \Delta F1_{FB}(n) \quad \text{(EQ 3)}$$

where λ_{FF} is a learning rate parameter for the feedforward command. In words, the feedforward command for the next trial is updated by adding some fraction (characterized by λ_{FF}) of the feedback-based corrective command for the current trial.

To fit the model to the data from the sham and anodal stimulation conditions, a particle swarm optimization procedure was used to find optimized values of the three free parameters of the model (α_A, α_S, and λ_{FF}) to fit the mean data for each block in each condition. The parameter estimates resulting from this procedure were highly robust to initial conditions, indicative of reaching the global minimum of the root mean square error (RMSE) measure.

Additionally, we examined the fit to our data of an alternative state-space model previously used to estimate and quantify learning and sensitivity to errors during motor learning (Thoroughman and Shadmehr 2000; Smith et al. 2006; Galea et al. 2015; Huberdeau et al. 2015). This model yielded qualitatively similar results, which are included in the Supplemental Materials (**Figure S3**).

4.2 DIVA model fits

The DIVA model fits to the two experimental conditions are provided in **Figure 4A**. In both cases, the model fit falls within the standard error of the sample mean for all blocks except the ramp block (block 20) and immediately after cessation of auditory feedback perturbation (sham: fit normalized RMSE = 0.01; Pearson's $r = 0.93$; anodal: fit normalized RMSE = 0.01; $r = 0.95$).

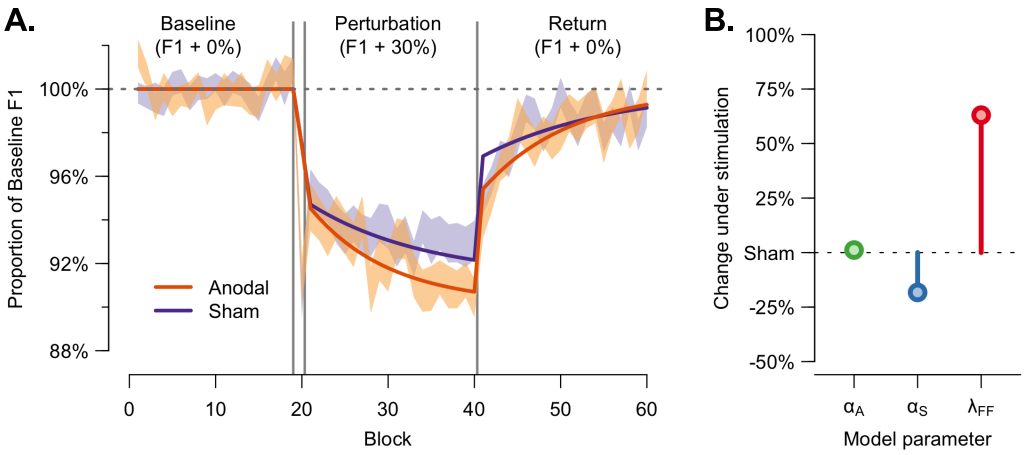

Figure 4. SimpleDIVA model fits to behavioral data.
(A) Solid lines depict the best-fit models identified by SimpleDIVA model simulations for both anodal tDCS (orange) and sham tDCS (purple). The shaded regions indicate the standard errors around the mean for the behavioral data, shown here for comparison with the models. **(B)** The percent change of our free parameter estimates is shown for anodal stimulation with respect to sham stimulation. The auditory gain factor (α_A) is shown in green, somatosensory gain factor (α_S) in blue, and the learning coefficient (λ_{FF}) is in red.

Table 1 compares the model parameter values for the two stimulation conditions. Whereas the values for the auditory feedback control gain, α_A, are nearly the same for the two conditions ($\alpha_A = 0.172$ during sham, $\alpha_A = 0.174$ during anodal stimulation, an increase of 1%), the somatosensory feedback control gain, α_S, decreased by 18% from $\alpha_S = 0.372$ in the sham condition to 0.304 in the anodal stimulation condition, and the value of the trial-to-trial feedforward command learning rate λ_{FF} increased by 63% in the anodal stimulation condition ($\lambda_{FF} = 0.194$) compared to the sham condition ($\lambda_{FF} = 0.119$). This is also shown graphically in **Figure 4B**. (To illustrate how the values of the free-parameters of this model affect the slope and magnitude of sensorimotor adaptation, three series of simulations in which only one parameter varies at a time are visualized in **Figure S4**.)

Parameter	Sham tDCS Estimate	Anodal tDCS Estimate	Difference
α_A	0.172	0.174	+1.15%
α_S	0.372	0.304	-18.28%
λ_{FF}	0.119	0.194	+63.03%

Table 1. SimpleDIVA model best fit parameter estimates.
Best fit parameter values for model simulations of the sham and anodal tDCS conditions. No differences between the two conditions was found for auditory feedback control gain (α_A), while somatosensory feedback control gain (α_S) decreased during anodal tDCS and feedforward learning/adaptation rate (λ_{FF}) increased under anodal tDCS relative to sham stimulation.

5. Discussion

The results of this study extend our understanding of the mechanisms through which speakers learn to adjust their feedforward motor plans in response to perturbed sensory feedback during speech production. When applying noninvasive neurostimulation over left ventral sensorimotor cortex, we observed an increased rate of adaptation responses to perturbed auditory feedback. Moreover, we found that this effect was specific to F1—the perturbed feature—and did not generalize to F2, indicating a task-specific effect rather than a global modulation of motor control processes.

The rate of increasing adaptive responses potentially depends on both the gain of the auditory feedback control subsystem for speech, which is responsible for within-trial reflexive responses to perceived auditory errors, and the rate of learning of feedforward commands, which is responsible for trial-to-trial increases in the anticipatory component of the compensatory response. Because the mechanisms by which tDCS affects cortical activity are uncertain and may be specific to a study's particular task (see Bortoletto et al., 2015), theoretically either error sensitivity, cortical plasticity, or both could have been modulated during anodal stimulation. We therefore utilized computational simulations using a simplified version of an established model of speech motor control, the DIVA model (Kearney et al., 2020), to decompose the adaptation responses into distinct, mechanistically precise components. Specifically, we extracted estimates of three key parameters characterizing the main control subsystems of the speech motor controller – the auditory feedback control gain, the somatosensory feedback control gain, and the feedforward command learning/adaptation rate – under anodal tDCS and sham stimulation. These simulations indicated that stimulation resulted in an increase in the feedforward learning rate, whereas the auditory feedback control gain was essentially unaffected by the perturbation. Thus, the performance gains resulting from stimulation were presumably due to increased adaptation of the feedforward commands for subsequent productions rather than increased within-trial reflexive responses by the auditory feedback controller.

Furthermore, best-fit models also included a small, unanticipated decrease in the gain of the somatosensory feedback control subsystem. While a change in somatosensory feedback control gain may initially be surprising given the auditory perturbation used in this study, a change in this parameter makes sense when its role is considered in context of the feedback control system in aggregate: Under normal circumstances, the somatosensory control subsystem *counteracts* compensatory adjustments to auditory perturbations, because these adjustments have the effect of producing somatosensory feedback that mismatches the somatosensory target for a given articulation. Teleologically, decreasing the gain of the somatosensory feedback control subsystem reduces this counteraction, allowing for more complete motor adaptation to the auditory perturbation, as seen in the larger magnitude of adaptation under anodal tDCS (**Figure 2B**). Mechanistically, however, the source of the change in the somatosensory feedback control gain is less certain. This change may reflect the relatively limited spatial resolution of tDCS, in that our electrode montage also likely resulted in stimulating current to left ventral somatosensory cortical areas in postcentral gyrus (**Figure 1D**) – including tissue comprising somatosensory state, target, and error maps (Guenther, 2016) – in addition to speech motor control areas in left vPMC and vMC. Given the spatial proximity of motor and somatosensory cortex, identifying the causal mechanisms that affect the integration of somatosensory information will require stimulation approaches with greater spatial specificity. For instance, noninvasive techniques such as TMS, and invasive techniques such as cortical cooling, have been used to make finer-grained functional dissociations between adjacent perisylvian neuroanatomy (e.g., Pulvermüller et al., 2006; Long et al., 2016), and applying computational modeling to the behavioral changes measured under more targeted stimulation may offer insight into mechanistic changes to sensory feedback gain control.

Previous work on sensorimotor adaptation during auditory feedback perturbation indicated that individual differences in participants' auditory acuity, or ability to detect feedback errors, explained a

significant portion of the variance in the degree of adaptation measured across subjects (Villacorta et al., 2007; Ghosh et al., 2010). Here, we show that it is not auditory error detection that increases under stimulation of ventral sensorimotor cortex, but rather the motor adaptation rate. This is not to say that auditory acuity does not play a role in adaptation, but that the areas we stimulated do not appear to mediate auditory acuity or error detection; instead, these areas must support, at some level, updating of stored motor programs for speech sounds, consistent with the DIVA model. The DIVA model also predicts that, in contrast with left vPMC, right vPMC is responsible for transforming auditory and somatosensory error signals into corrective motor commands. We therefore hypothesize that if the right vPMC were stimulated using anodal tDCS, we would see modulation of auditory and somatosensory feedback control gains but not the feedforward command adaptation rate. Testing these predictions must be the goal of future work that compares differences in reflexive vs. adaptive responses under right- vs. left-hemisphere stimulation.

It is important to note that while the results of this study are consistent with a model in which left sensorimotor cortex plays a causal role in sensorimotor adaptation, the evidence presented here is by itself not sufficient to establish the unique causal involvement of this region, as we did not test for effects of anodal stimulation on some other putatively unrelated control region. It may be the case that anodal stimulation to any region of brain increases the rate learning. However, while both empirical (Huang et al., 2017) and modelling (Datta et al., 2013) work on the physiological effects of targeted transcranial electrical stimulation suggest that these can be quite focal depending on local field strength, further work remains necessary to establish the causal contribution of this or other cortical areas in sensorimotor adaptation.

Further support for the view that anodal tDCS of left sensorimotor cortex affects learning rate and not error detection is the lack of differences between anodal and sham stimulation on compensation magnitude during the first perturbation trial. If anodal stimulation enhanced error detection during compensation, we might have observed those differences earlier in the perturbation phase of the session, during a period in which compensation should be dominated by reflexive responses. A previous study employing anodal tDCS over either motor cortex or cerebellum during visuomotor adaptation by Galea and colleagues (2011) showed early effects on compensation when stimulation was applied to cerebellum, but significant after-effects when stimulation was applied to motor cortex, providing further evidence that motor cortex supports adaptive responses. Lametti and colleagues (2017) also reported distinct roles for motor cortex and cerebellum in a similar auditory feedback perturbation study where tDCS was applied to either brain region. In order to better understand the extent of the dissociation between the neural mechanisms supporting adaptive and reflexive responses, future studies are needed to directly compare the effects of tDCS on sensorimotor adaptation to those during unexpected/random feedback perturbations, in which adaptive, but not reflexive, responses should be reduced. Furthermore, studies directly comparing reflexive and adaptive processes can be used to inform modifications to the simplified DIVA model to better capture participants' behavior during transitional periods during the present paradigm, given that our model was least successful estimating behavior during the ramp and beginning of the return phase.

We observed a significant effect of anodal tDCS on F1 trial-to-trial variability compared to sham. Because we did not find any interaction between stimulation condition and phase of each session, we might conclude that anodal tDCS caused increased variability that was not related to the increased feedforward learning rate. However, the fact that we did not see a corresponding difference in F2 production variability suggests instead that the effects on F1 *are* related to learning, and perhaps the differences with phase of the experiment are too small to observe at our current power. We performed post-hoc paired Student's t-tests on the degree of variability during anodal and sham tDCS for each phase and observed that, whereas the baseline phase showed little difference between stimulation conditions (two-tailed; $t_{17} = 0.57$, $p = 0.58$, Cohen's $d = 0.20$), the perturbation and return phases showed differences that trended in the direction of higher variability during anodal tDCS than sham (perturbation: $t_{17} = 2.03$, $p = 0.06$, $d = 0.54$; return: $t_{17} = 1.92$, $p = 0.07$, $d = 0.67$). Before we can draw conclusions about how

variability and learning might be related in this paradigm, we may need to better understand speech motor variability under tDCS without auditory feedback perturbations, which is as of yet, unstudied.

In addition to the sensorimotor adaptation we observed during the perturbation phase of each session, we also recorded a significant downward shift in baseline F1 after the first session, despite requiring a minimum of 7 days between visits to the lab. Previous work from Heald and Nusbaum (2015) found remarkable day-to-day consistency in the acoustics of speakers' vowel productions and so our observation is notable, especially given that the perturbation periods lasted only approximately five minutes per session. We did not find any interaction between session number and stimulation condition; therefore, we do not have evidence that increased adaptation during anodal tDCS had any long-term effects on speech production beyond repeating the behavioral task. This raises the possibility that behavioral interventions with repeated sensory feedback perturbation may be useful in training or retraining target outputs in speech motor learning, such as in the case of second language learning, vocal accent or gender modification, or speech motor recovery following pathology. However, we also noted a reduced rate of adaptation in the third session compared to the second. While this session-by-time effect did not interact with stimulation manipulation, it indicates that participants' susceptibility to the perturbation manipulation itself may differ when undergoing repeated treatments. Whether this is due to the accumulation of changes observed in baseline speech production targets measured at the third session, decreased sensitivity to auditory perturbations, or some other factor remains a question for future research. Indeed, before we can assess the applied or clinical utility of such paradigms, future work will need to assess individual consistency in rate and magnitude of speech motor adaptation over multiple sessions, without involving tDCS. This will also provide important information about how we might control for session-to-session learning in within-subjects experimental designs more completely.

Approximately two thirds of participants who completed the screening session of our study showed significant adaptation during the auditory feedback perturbation. We chose to focus the stimulation aspect of the study on those who showed adaptation because variability in this behavior has been documented (e.g., Purcell and Munhall, 2006) but has yet to be successfully explained. There are several hypotheses as to why certain people do not adapt, such as inability to perceive the perturbation due to poor auditory acuity (Ghosh et al., 2010) or a stronger adherence to somatosensory speech targets than auditory ones. It may be useful to perform future studies with these participants to determine their auditory acuity and to see how their behavior is affected by tDCS. Given the results of our model simulations, it is possible that anodal tDCS to left ventral sensorimotor cortex could cause non-adapting participants to depend less on their somatosensory feedback and therefore increase the magnitude of their adaptation responses. However, even within our adapting participants, we observed individual differences in behavior. Given that some individuals adapt more than others, we tested for correlations in adaptation magnitude across conditions and sessions. We found a significant correlation between sessions 1 and 2, and not between 1 and 3; however, the numerical difference between these two correlations was small, and so we do not feel we have sufficient evidence to say whether or not individual differences in overall adaptation magnitude are demonstrated by our results. Though most of our participants showed increased adaptation during anodal stimulation relative to the sham condition, few showed either no effect of tDCS or an opposite pattern of behavior. Some of these differences may be attributed to high variability in speech production across trials and sessions, but we cannot rule out the possibility that tDCS affects some people differently (e.g., Schaal et al., 2015). As a relatively new technology, more work is needed to better understand sources of behavioral variability under tDCS.

In summary, participants showed increased sensorimotor adaptation under anodal tDCS to left ventral sensorimotor cortex during perturbed auditory feedback, demonstrating the ability of noninvasive brain stimulation to enhance how speakers learn to integrate sensory feedforward and feedback speech motor commands to modify stored motor programs for speech. Through computational modeling, we were able to verify the effects of anodal tDCS on sensorimotor learning and gain insights into the cortical

mechanisms that limit adaptation to ongoing perceived auditory errors. The results of this study further our knowledge of the cortical mechanisms supporting the speech motor system's ability to adapt in response to altered sensory feedback. Additionally, these findings have implications for understanding how to effectively deploy tDCS as both a research instrument and a therapeutic technique in treatment of speech motor control issues involving abnormal feedback-based adaptation, such as stuttering (Cai et al., 2012; Daliri et al., 2017; Chesters et al., 2018), aphasia (Behroozmand et al., 2018), and Parkinson's disease (Abur et al. 2018).

6. Acknowledgments

We thank Jason Tourville, Alfonso Nieto-Castañón, Shanqing Cai, Sara Dougherty, Jennifer Golditch, Elly Hu, Cecilia Cheng, Emily Thurston, and Ja Young Choi. This research was supported by NIDCD of the NIH under award numbers R03DC014045 to TP, and R01DC002852 to FG. TLS was supported by T90DA032484. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

7. References

Abur DA, Lester-Smith RA, Daliri A, Lupiani AA, Guenther FH, Stepp CE (2018) Sensorimotor adaptation of voice fundamental frequency in Parkinson's disease. PLoS One 13: e0191839.

Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmation hypothesis testing: Keep it maximal. J Mem Lang 68:255-278.

Basilakos A, Smith KG, Fillmore P, Fridriksson J, Fedorenko E (2018) Functional characterization of the human speech articulation network. Cereb Cortex 28:1816-1830.

Bates D, Mächler M, Bolker B, Walker S (2014) Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

Behroozmand R, Phillip L, Johari K, Bonilha L, Rorden C, Hickok G, Fridriksson J (2018) Sensorimotor impairment of speech auditory feedback processing in aphasia. Neuroimage 165:102-111.

Behroozmand R, Shebek R, Hansen DR, Hiroyuki O, Robin DA, Howard MA, III, Greenlee JDW (2015) Sensory-motor networks involved in speech production and motor control: An fMRI study. Neuroimage 109:418-428.

Boersma P (2001) Praat, a system for doing phonetics by computer. Glot International 5:341-345.

Bortoletto M, Pellicciari MC, Rodella C, Miniussi C (2015) The interaction with task-induced activity is more important than polarization: A tDCS study. Brain Stimul 8:269-276.

Brunoni AR, Amadera J, Berbel B, Volz MS, Rizzerio BG, Frengi F (2011) A systematic review on reporting and assessment of adverse effects associated with transcranial direct current stimulation. International Journal of Neuropsychopharmacology 14:1133-1145.

Buchwald A, Calhoun H, Rimikis S, Lowe MS, Wellner R, Edwards DJ (2019) Using tDCS to facilitate motor learning in speech production: The role of timing. Cortex 111:274-285.

Burnett TA, Freedland MB, Larson CR, Hain TC (1998) Voice f0 responses to manipulations in pitch feedback. J Acoust Soc Am 103:3153-3161.

Cai S, Boucek M, Ghosh SS, Guenther FH, Perkell JS (2008) A system for online dynamic perturbation of formant trajectories and results from perturbations of the Mandarin triphthong /iau/. Presented at the Eighth International Seminar on Speech Production, Strasbourg, France, December 8-12.

Cai S, Beal DS, Ghosh SS, Tiede MK, Guenther FH, Perkell JS (2012) Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. PLoS One 7:e41830.

Chen SH, Lui H, Xu Y, Larson CR (2007) Voice F0 responses to pitch-shifted voice feedback during English speech. J Acoust Soc Am 121:1157-1163.

Chesters J, Möttönen R, Watkins KE (2018) Transcranial direct current stimulation over left

inferior frontal cortex improves speech fluency in adults who stutter. Brain 141:1161-1171.

Daliri A, Weiland EA, Cai S, Guenther FH, Chang SE (2017) Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter. Dev Sci 21:e12521.

Datta A, Bansal V, Diaz J, Patel J, Reato D, Bikson M (2009) Gyri-pricise head model of transcranial direct current stimulation: Improved spatial focality using a ring electrode versus conventional rectangular pad. Brain Stimul 2:201-207.

Datta A, Zhou X, Su Y, Parra LC, Bikson M (2013) Validation of finite element model of transcranial electrical stimulation using scalp potentials: implications for clinical dose. J Neural Eng 10:036018.

Dayan E, Censor N, Buch ER, Sandrini M, Cohen LG (2013) Noninvasive brain stimulation: from physiology to network dynamics and back. Nat Neurosci 16:838-844.

Diedenhofen B, Musch J (2015) cocor: A comprehensive solution for the statistical comparison of correlations. PLoS ONE, 10(4): e0121945.

Fertonani A, Rosini S, Cotelli M, Rossini PM, Miniussi C (2010) Naming facilitation induced by transcranial direct current stimulation. Behav Brain Res 208:311-318.

Filmer HL, Dux PE, Mattingley JB (2014) Applications of transcranial direct current stimulation for understanding brain function. Trends Neurosci 37:742-753.

Galea JM, Mallia E, Rothwell J, Diedrichsen J (2015) The dissociable effects of punishment and reward on motor learning. Nature Neurosci 18:597-602.

Galea JM, Vazquez A, Pasricha N, Orban de Xivry JJ, Celnik P (2011) Dissociating the roles of the cerebellum and motor cortex during adaptive learning: The motor cortex retains what the cerebellum learns. Cereb Cort 21:1761-1770.

Ghosh SS, Tourville JA, Guenther FH (2008) A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. J Speech Lang Hear Res 51:1183-1202.

Ghosh SS, Matthies ML, Maas E, Hanson A, Tiede M, Ménard L, Guenther FH, Lane H, Perkell JS (2010) An investigation of the relation between sibilant production and somatosensory and auditory acuity. J Acoust Soc Am 128:3079-3087.

Golfinopoulos E, Tourville JA, Guenther FH (2010) The intergration of large-scale neural network modeling and functional brain imaging in speech motor control. Neuroimage 52:862-874.

Guenther FH (1994) A neural network model of speech acquisition and motor equivalent speech production. Biol Cybern 72:43-53.

Guenther FH (1995) Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. Psychol Rev 102:694-621.

Guenther FH (2016) Auditory Feedback Control. In: Neural control of Speech, pp153-176. Cambridge, MA: The MIT Press.

Guenther FH, Ghosh SS, Tourville JA (2006) Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang 96:280-301.

Hain TC, Burnett TA, Kiran S, Larson CR, Singh S, Kenney MK (2000) Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. 130:133-141.

Houde JF, Jordan MI (1998) Sensorimotor adaptation in speech production. Science 279:1213-1216.

Huang Y, Liu AA, Lafon B, Friedman D, Dayan M, Wang X, Bikson M, Doyle WK, Devinsky O, Parra LC (2017) Measurements and models of electric fields in the *in vivo* human brain during transcranial electric stimulation. eLife 6:e18834.

Huberdeau DM, Krakauer JW, Haith AM (2015) Dual-process decomposition in human sensorimotor adaptation. Curr Opin Neurobiol 33:71-77.

Jones JA, Munhall KG (2000) Perceptual calibration of *F*0 production: Evidence from feedback perturbation. J Acoust Soc Am 108:1246-1251.

Katseff S, Houde J, Johnson K (2011) Partial compensation for altered auditory feedback: A tradeoff with somatosensory feedback? Lang Speech 55:295-308.

Kearney E, Nieto-Castañón A, Weerathunge HR, Falsini R, Daliri A, Abur D, Ballard K, Chao S, Heller Murray L, Scott TL, Guenther F (2020) A simple 3-parameter model for examining adaptation in speech and voice production. Front Psychol. doi: 10.3389/fpsyg.2019.02995

Kuo HI, Bikson M, Datta A, Minhas P, Paulus W, Kuo MF, Nitsche MA (2013) Comparing cortical plasticity induced by conventional and high-

definition 4 × 1 ring tDCS: A neurophysiological study. Brain Stimul 6:644-648.

Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest package: tests in linear mixed effects models. Journal of Statistical Software, 82(13).

Lametti DR, Nasir SM, Ostry DJ (2012) Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. J Neurosci 32:9351-9358.

Lametti DR, Smith HJ, Freidin PF, Watkins KE (2017) Cortico-cerebellar networks drive sensorimotor learning in speech. J Cog Neurosci 30:540-551.

Lawrence, MA (2013) ez: Easy Analysis and Visualization of Factorial Experiments. R package version 4.3.

Long MA, Katlowitz KA, Svirsky MA, Clary RC, McAllister T, Majaj N, Oya H, Howard MA III, Greenlee JDW (2016) Functional segregation of cortical regions underlying speech timing and articulation. Neuron 89:1187-1193.

Malyutina S, Den Ouden DB (2015) High-definition tDCS of noun and verb retrieval in naming and lexical decision. NeuroRegul 2:111-125.

Meteyard L, Davies RAI (2020) Best practice guidance for linear mixed-effects models in psychological science. J Mem Lang 112:104092.

Monti A, Ferrucci R, Fumagalli M, Mameli F, Cogiamanian F, Ardolino G, Priori A (2013) Transcranial direct current stimulation (tDCS) and language. J Neurol Neurosurg Psychiatry 84:832-842.

Nasir SM, Ostry DJ (2009) Auditory plasticity and speech motor learning. Proc Natl Acad Sci 106:20470-20475.

Nitsche MA, Paulus W (2000) Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. J Physiol 527:633-639.

Nitsche MA, Roth A, Min-Fang K, Fischer AJ, Liebetanz D, Lang N, Tergau F, Paulus W (2007) Timing dependent modulation of associative plasticity by general network excitability in the human motor cortex. J Neurosci 27:3807-3812.

Niziolek CA, Guenther FH (2013) Vowel category boundaries enhance cortical and behavioral responses to speech feedback. J Neurosci 33:12090-12098.

Pulvermüller F, Huss M, Kherif F, del Prado Martin FM, Hauk O, Shtyrov Y (2006) Motor cortex maps articulatory features of speech sounds. Proc Natl Acad Sci USA 103:7865-7870.

Purcell DW, Munhall KG (2006) Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. J Acoust Soc Am 120:966-977.

Redford MA (2019) Speech production from a developmental perspective. J Speech Lang Hear Res 62:2946-2962.

Rroji O, van Kuyck K, Nuttin B, Wenderoth N (2015) Anodal tDCS over the primary motor cortex facilitates long-term memory formation reflecting use-dependent plasticity. PLoS One 10:e0127270.

Schall NK, Krause V, Lange K, Banissy MJ, Williamson VJ, Pollock B (2015) Pitch memory in nonmusicians and musicians: Revealing functional differences using transcranial direct current stimulation. Cereb Cort 25:2774-2782.

Scott SH (2004) Optimal feedback control and the neural basis of volitional motor control. Nat Rev Neurosci 5:532-546.

Shadmehr R, Smith MA, Krakauer JW (2010) Error correction, sensory prediction, and adaptation in motor control. Annu Rev Neurosci 33:89-108.

Shiller DM, Sato M, Gracco VL, Baum SR (2009) Perceptual recalibration of speech sounds following speech motor learning. J Acoust Soc Am 125:1103-1113.

Smith MA, Ghazizadeh A, Shadmehr R (2006) Interacting adaptive processes with different timescales underlie short-term motor learning. PLoS Biol 4:e179.

Thoroughman KA, Shadmehr R (2000) Learning of action through adaptive combination of motor primatives. Nature 407:742-747.

Tourville JA, Reilly KJ, Guenther FH (2008) Neural mechanisms underlying auditory feedback control of speech. Neuroimage 39:1429-1443.

Tourville JA, Guenther FH (2011) The DIVA model: A neural theory of speech acquisition and production. Lang Cogn Process 26:952-981.

Villacorta VA, Perkell JS, Guenther FH (2007) Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. J Acoust Soc Am 122:2306-2319

SUPPLEMENTAL MATERIALS

Table S1: Adaptation of F1 during perturbation phase (see §3.2.1)

Fixed effect term*	Estimate (β)**	Std. Err.	df	*t*-value	*p*-value
(intercept)	0.93	0.0010	17.49	91.66	$\ll 0.0001$
Stimulation	0.0061	0.0058	16.00	1.06	0.31
Time	-0.00049	0.00011	16.17	-4.65	0.00026
Session	-0.0062	0.0058	16.00	-1.07	0.30
Stimulation × Time	0.00018	0.000060	2011.00	2.99	0.0028
Stimulation × Session	-0.0032	0.0098	16.00	-0.32	0.75
Time × Session	-0.00014	0.000060	2011.00	-2.40	0.017
Stimulation × Time × Session	-0.00016	0.00011	16.17	-1.52	0.15

*Model: *ΔF1 ~ Stimulation * Time * Session + (1 + Stimulation + Time + Session | Participant) + (1 | Word)*
**Based on deviation (sum) coded contrasts.

Table S2: Adaptation of F2 during perturbation phase (see §3.2.1)

Fixed effect term*	Estimate (β)**	Std. Err.	df	*t*-value	*p*-value
(intercept)	1.00	0.0036	14.48	281.02	$\ll 0.0001$
Stimulation	-0.00070	0.0027	15.96	-0.26	0.80
Time	0.000016	0.000075	16.56	0.22	0.83
Session	0.0041	0.0027	15.96	1.53	0.15
Stimulation × Time	-0.000049	0.000046	2014.00	-1.08	0.28
Stimulation × Session	-0.0020	0.0032	16.00	-0.61	0.55
Time × Session	0.000089	0.000046	2014.00	1.96	0.05
Stimulation × Time × Session	-0.000043	0.000075	16.56	-0.57	0.57

*Model: *ΔF2 ~ Stimulation * Time * Session + (1 + Stimulation + Time + Session | Participant) + (1 | Word)*
**Based on deviation (sum) coded contrasts.

Table S3: Recovery of F1 during return phase (see §3.2.2)

Fixed effect term*	Estimate (β)**	Std. Err.	df	*t*-value	*p*-value
(intercept)	0.98	0.0067	9.93	146.41	$\ll 0.0001$
Stimulation	0.0012	0.0060	16.71	0.19	0.85
Time	0.00057	0.000099	15.80	5.82	$\ll 0.0001$
Session	-0.0055	0.0060	16.71	-0.92	0.37
Stimulation × Time	-0.000090	0.000055	2032.00	-1.63	0.11
Stimulation × Session	-0.0018	0.0053	16.00	-0.34	0.74
Time × Session	0.000090	0.000055	2032.00	1.64	0.10
Stimulation × Time × Session	0.000084	0.000099	15.80	0.86	0.41

*Model: *ΔF1 ~ Stimulation * Time * Session + (1 + Stimulation + Time + Session | Participant) + (1 | Word)*
**Based on deviation (sum) coded contrasts.

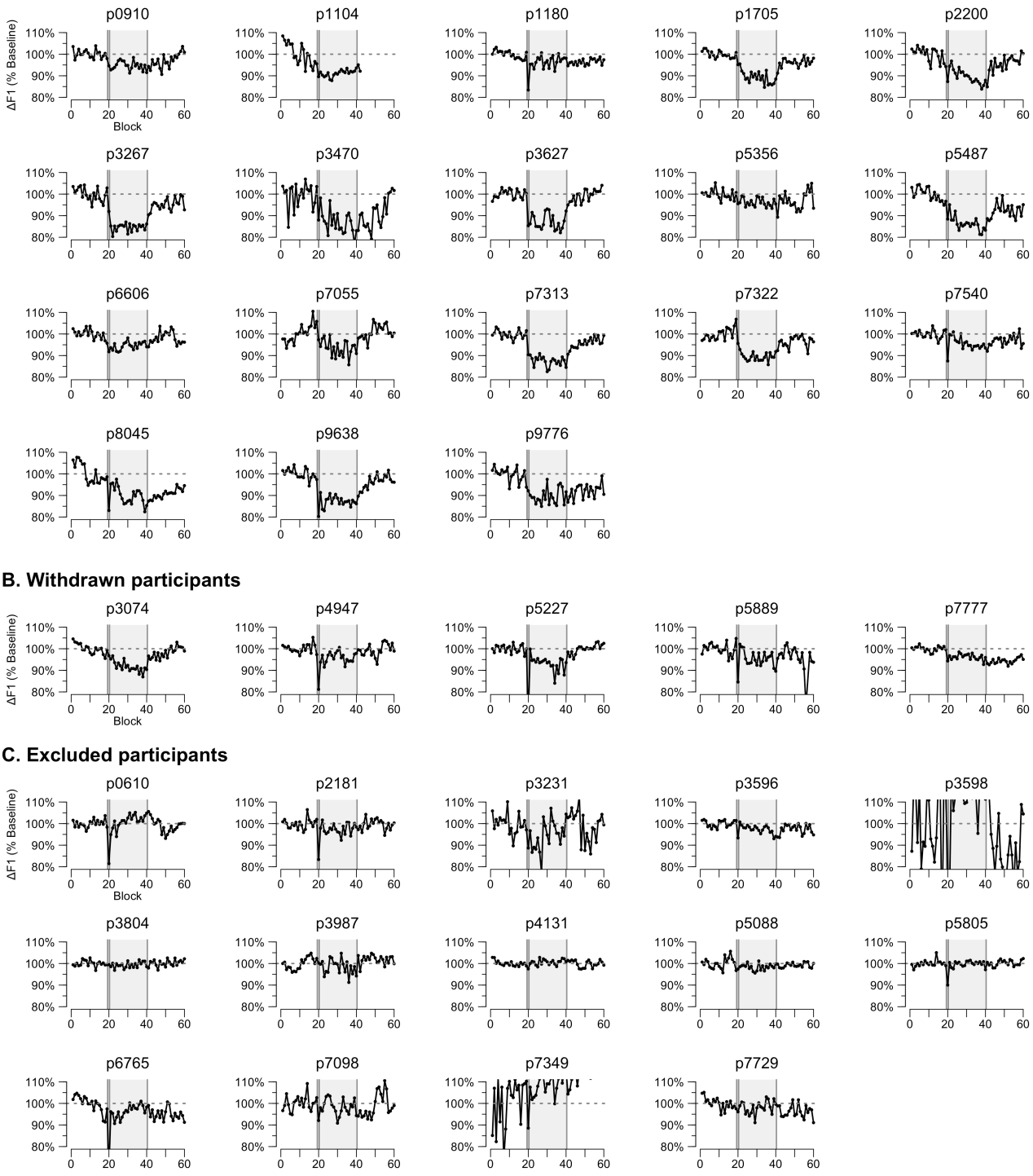

Figure S1. Screening session data from individual participants.
Participants' F1 productions were averaged within blocks of three trials for display; all statistical tests were performed on single trial data. The shaded region indicates the perturbation phase of the experiment. Each panel corresponds to a single participants' data. The double vertical gray line indicates the beginning and end of the "ramp" phase. **(A)** Participants who showed significant adaptation as defined by our inclusion criterion and completed all tDCS sessions ($N = 18$). Data loss occurred during the return phase for participant "p1104" (top row, second column). **(B)** Participants who met our inclusion criterion for adaptation but did not return to the lab to complete subsequent tDCS sessions. **(C)** Participants who did not meet our inclusion criterion. Of these participants, two showed "following" responses; adaptive responses that do not oppose the perturbation but mirror its direction. These participants are "p0610" (first row, first column) and "p7349" (third row, third column).

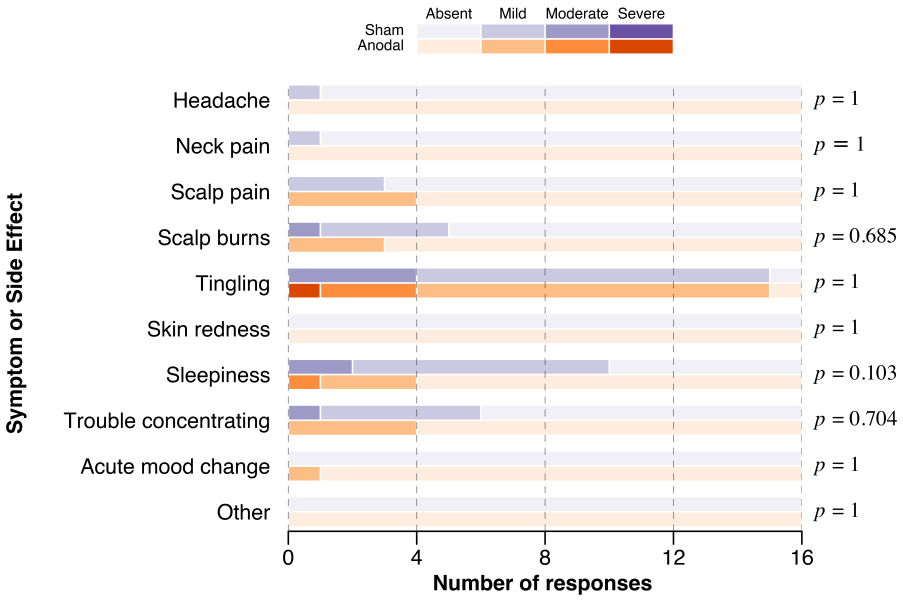

Figure S2. Incidence of self-reported symptoms or side effects of HD-tDCS.
After each HD-tDCS session, participants completed the survey described by Bruononi et al. (2011) to gauge the presence and degree of any adverse experiences with HD-tDCS. The frequency of each sensation and their magnitudes are shown for sham (top, purple) and anodal (bottom, orange) stimulation conditions. Participants' most frequently reported experiencing mild scalp tingling. At right, *p*-values indicate whether participants' propensity to report the severity of these sensations differed between anodal and sham stimulation, computed using Fisher's exact test. Participants were no more likely to report the presence or magnitude of any adverse experience with HD-tDCS in the anodal condition compared to sham stimulation, suggesting participants were effectively blinded to the stimulation they were receiving. Post-stimulation questionnaires were not recorded for two participants.

Methods for state-space models

In the main text, we report computational simulations using a simplified version of the DIVA model to better interpret the mechanisms by which anodal tDCS affected sensorimotor adaptation during our auditory perturbation task. Here, we report additional simulations with another class of models in order to compare results. Previous studies of limb motor learning have used state-space models to estimate and quantify learning and sensitivity to errors (Thoroughman and Shadmehr 2000; Smith et al. 2006; Galea et al. 2015; Huberdeau et al. 2015). In these models, it is assumed that the central nervous system (CNS) learns from errors; based on its prediction of the next state of the perturbation, the CNS then produces a movement that compensates for the predicted perturbation. In other words, the learner updates his/her estimate of the magnitude of perturbations in each trial using his/her previous estimate of perturbations and the magnitude of error experienced in the current trial.

Here, we used a similar procedure to estimate learning across trials. Although several state-space models (e.g., dual rate vs. single rate) have been suggested for motor adaptation (Smith et al. 2006), for simplicity we used a single rate state-space model as follows:

$$\Delta F1_{PerceivedError}(n) = \Delta F1_{Perturbation}(n) - \Delta F1_{Correction}(n-1) \quad \text{(EQ S1)}$$

$$\Delta F1_{Correction}(n) = \alpha \times \Delta F1_{Correction}(n-1) + \beta \times \Delta F1_{PerceivedError}(n) \quad \text{(EQ S2)}$$

$$F1_{Produced}(n) = F1_{Baseline} - \Delta F1_{Correction}(n) \quad \text{(EQ S3)}$$

Equation S1 defines the perceived auditory error in F1 on the current trial ($\Delta F1_{PerceivedError}(n)$) as the difference between the F1 perturbation on that trial and the F1 correction estimate from the previous trial.

The F1 correction estimate for the current trial given in **Equation S2** is described as a weighted sum of the correction estimate from the previous trial and the perceived error on the current trial. The parameter α describes the weight given to the correction estimate from the previous trial (i.e. how much of the correction is "remembered") and β is a parameter which describes the sensitivity to the current perceived auditory error.

We fit this state-space model using the same particle swarm optimization procedure as in our DIVA model simulations. Here we computed optimized values of two free parameters (α and β). Measures of model fit for the state-space models were comparable to those achieved by the DIVA model for both sham (fit normalized RMSE = 0.01, Pearson's r = 0.91; compare to RMSE = 0.01 and r = 0.93 for DIVA) and anodal stimulation conditions (fit normalized RMSE = 0.01, Pearson's r = 0.94; compare to RMSE = 0.01 and r = 0.95 for DIVA). **Figure S3** illustrates the best-fit state-space models compared to the participant behavioral data.

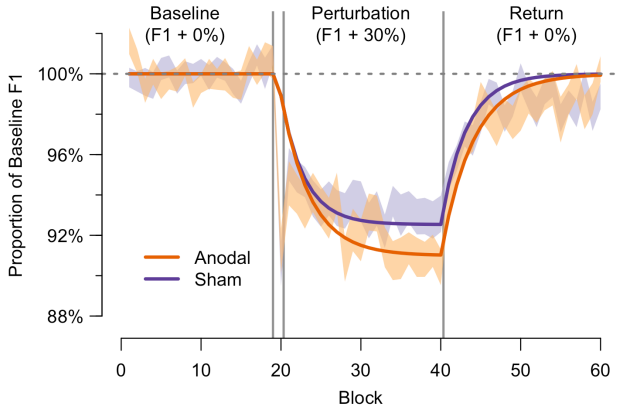

Figure S3: State-space model fits to participant formant data.
Solid lines depict the best-fit models identified by state-space model simulations for both anodal tDCS (orange) and sham tDCS (purple). The shaded regions indicate the standard errors around the mean for the behavioral data, shown here for comparison with the models. The double vertical gray line indicates the beginning and end of the "ramp" phase.

We compared the best-fit state-space models and found that α increased by 6.32% from $\alpha = 0.803$ during sham stimulation to $\alpha = 0.854$ during anodal stimulation. The β parameter decreased by 0.69% from $\beta = 0.0724$ during sham to $\beta = 0.0719$ during anodal stimulation. Though the changes in parameter values are small compared to those computed in the DIVA model simulations, we would like to note that the relatively larger parameter change between sham and anodal stimulation involves the parameter α. During anodal stimulation α is increased, illustrating a greater reliance on information from the previous trial in order to inform the corrective command during the current trial. This result is qualitatively similar to our interpretation of the DIVA model simulations in which the greatest factor affected during anodal stimulation was the learning parameter λ_{FF}, which also illustrates a greater reliance on information from the previous trial.

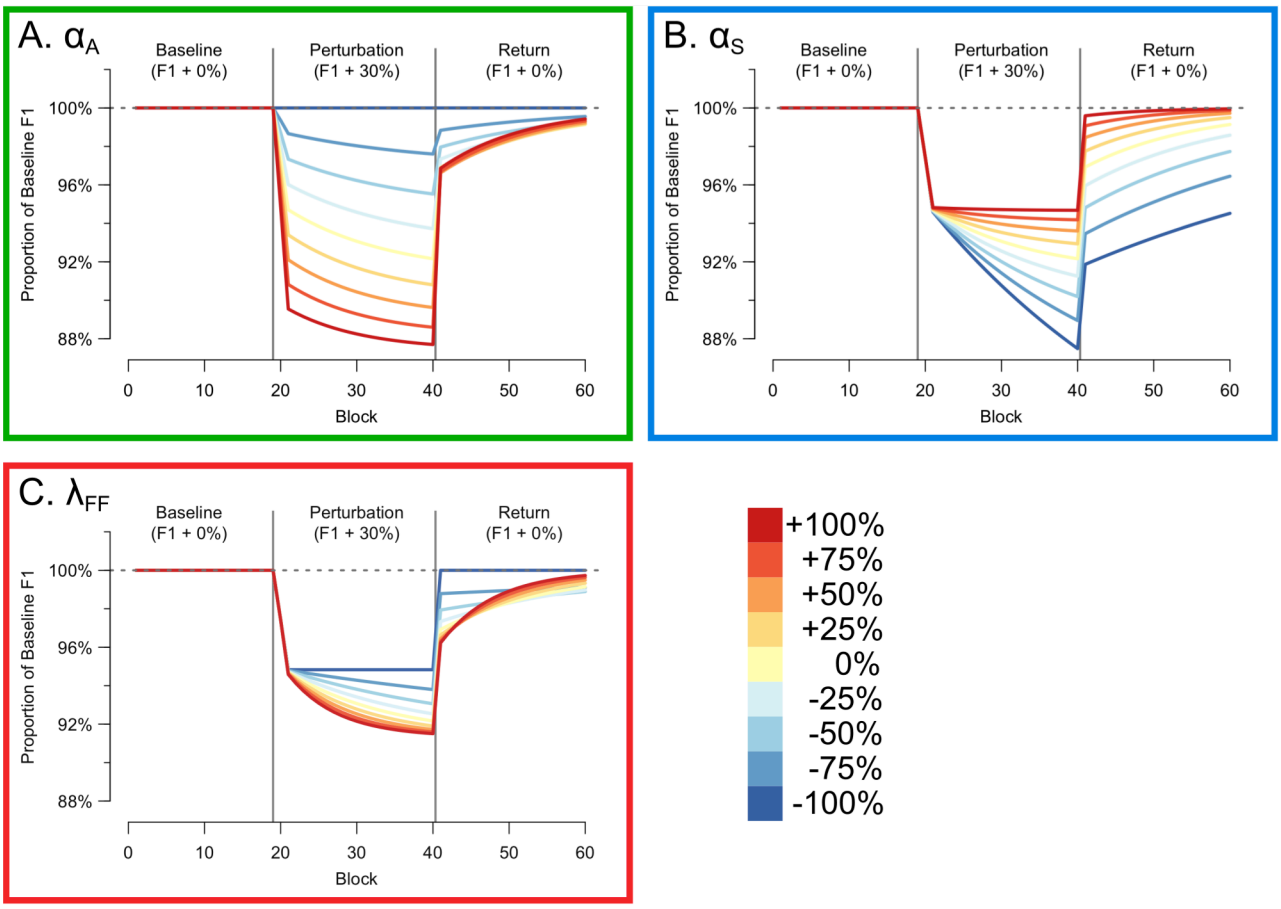

Figure S4: Simulated model parameter manipulations.
We simulated F1 percent change from baseline in our experiment while only allowing one of the model's free-parameters deviate at a time, with respect to its best fit value for sham stimulation. Warm colors indicate the effect of an increase in the parameter value, while cool colors indicate a decrease. **(A)** shows simulations with different values of auditory feedback control gain (α_A), **(B)** shows simulations with different values of somatosensory feedback control gain (α_S), and **(C)** shows simulations with different values of the feedforward command learning rate (λ_{FF}).